

<b>APPELLANT'S BRIEF UNDER 37 CFR §41.37</b>		Docket No. C0012/7000
Applicant:	Raju C. Bopardikar, Jack J. Stiffler, Jacob Y. Bast, Gary A. Cardone, David E. Kaufman, Stuart P. MacEachern, Bruce D. McLeod, James M. Nolan, Jr., Zdenek Radouch and James A. Wentworth, III	
Serial No:	09/608,521	
Filed:	June 30, 2000	
For:	METHOD AND APPARATUS FOR IMPLEMENTING HIGH-PERFORMANCE, SCALEABLE DATA PROCESSING AND STORAGE SYSTEMS	
Examiner:	A. M. Mirza	
Art Unit:	2145	

Mail Stop Appeal Brief - Patents  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

This brief is in furtherance of the Notice of Appeal, filed in this case on April 30, 2009.

The fees required under §41.20(b)(2), and any required petition for extension of time for filing this brief and fees therefor, are dealt with in the accompanying TRANSMITTAL OF APPEAL BRIEF.

A single copy of this brief is transmitted (37 C.F.R. §41.37(a)) and contains these items under the following headings, and in the order set forth below (37 C.F.R. §41.37(c)(1)):

- I REAL PARTY IN INTEREST
- II RELATED APPEALS AND INTERFERENCES
- III STATUS OF CLAIMS
- IV STATUS OF AMENDMENTS
- V SUMMARY OF CLAIMED SUBJECT MATTER
- VI GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL
- VII ARGUMENT
- VIII APPENDIX OF CLAIMS INVOLVED IN THE APPEAL
- IX EVIDENCE APPENDIX
- X RELATED PROCEEDINGS APPENDIX

#### **I REAL PARTY IN INTEREST (37 C.F.R. §41.37(c)(1)(i))**

The real party in interest in this appeal is EMC Corporation, 176 South Street, Hopkinton, MA 01748.

## **II RELATED APPEALS AND INTERFERENCES (37 C.F.R. §41.37(c)(1)(ii))**

There are no other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in the pending appeal.

## **III STATUS OF CLAIMS (37 C.F.R. §41.37(c)(1)(iii))**

### **A. TOTAL NUMBER OF CLAIMS IN APPLICATION**

The claims in the application are: 1-86.

### **B. STATUS OF ALL THE CLAIMS IN APPLICATION**

1. Claims pending: 1-11 and 44-54.
2. Claims canceled: None.
3. Claims withdrawn from consideration, but not canceled: 12-43 and 55-86.
4. Claims allowed: None.
5. Claims rejected: 1-11 and 44-54.

### **C. CLAIMS ON APPEAL**

The claims on appeal are: 1-11 and 44-54.

## **IV STATUS OF AMENDMENTS (37 C.F.R. §41.37(c)(1)(iv))**

No amendments have been filed.

## **V SUMMARY OF CLAIMED SUBJECT MATTER (37 C.F.R. §41.37(c)(1)(v))**

The present invention relates to a method and apparatus for providing storage services to clients from a "pool" of storage devices (see page 4, line 21 - page 5, line 13 and Figure 1). The clients (102, 104) are connected to access interfaces (106, 108, 122) which are, in turn, connected to the pool of storage devices (118, 120, 122, 124) by a switch fabric (114, 116). The access interfaces communicate with the devices in the storage device pool and, based on the workload of each device in the storage device pool, select a subset of the storage devices to use for any given transaction and distribute the workload (page 5, lines 16-22). Thus, the data received from a client is transferred to the selected subset of storage devices based on the workload of the devices rather than on a physical device address (page 5, line 23 – page 6, line 6). The

result of this operation is that data may be stored in a storage device other than the storage device from which that data was retrieved (page 7, line 22 – page 8, line 11).

The correspondence between data identification and the location at which the data is stored is broken by using data “handles”. In particular, in the present system, a client wishing to retrieve or store data specifies a data object (a set of data) and a data identifier or “handle” for that object instead of a traditional physical address that is associated with a location (page 10, lines 11-25). In the case of a read operation, this handle is used to locate a map (called an “fmap”) which specifies the actual physical locations of the data (page 10, lines 16-25). In the case of a write operation, the data is stored in physical addresses specified by allocation units which are preallocated to each interface module by the storage devices before any write operation takes place (page 12, lines 11-18). The allocation units which are used for a particular data storage operation are selected based on the recent activity of the storage device with which they are associated (page 14, lines 12-18). After the write operation is committed, the fmaps associated with the data handle are updated to reflect the new location of the data (page 13, lines 11-14).

Therefore, during a typical read-modify-write operation, data corresponding a data identifier would be read from a storage location and modified, but might be stored back into a location different from which it was read depending on resource load and availability. The data thus dynamically moves based on criteria determined by the data storage system. Consequently, load balancing becomes automatic (page 14, line 19- page 15, line 2).

The independent claims correspond to the specification as follows:

Claim 1 - an access interface module (page 5, lines 9-14):

that receives data storage requests, each including a data identifier, and data to be stored from the client (page 10, lines 11-25)

in response to each service request and based on a workload instead of a physical location in the plurality of devices, dynamically selects a subset of the plurality of storage devices in which the data is stored (page 5, lines 16-22 and page 14, lines 12-18, page 18, lines 12-24)

so that the physical storage device location to which data corresponding to the same data identifier is transferred can change from request to request in order to dynamically distribute the workload across the plurality of storage devices (page 14, line 19-page 15, line 2).

a switch fabric for temporarily connecting the access interface module to the selected subset of the plurality of storage devices so that the data can be transferred to the selected subset of storage devices (page 4, line 25 – page 5, line 8).

Claim 44

step (a) - page 10, lines 11-25

step (b) - page 5, lines 16-22 and page 14, lines 12-18; page 14, line 19-  
page 15, line 2

step (c) - page 4, line 25 – page 5, line 8

The dependent claims correspond to the specification as follows:

Claim 2 - page 14, lines 13-15

Claim 3 – page 4, line 25 – page 5, line 1, Figure 9, 952 and 942.

Claim 4 - page 5, lines 1-8.

Claim 5 – page 10, lines 4-10 Figure 3, 302, 304.

Claim 6 – page 22, lines 21-28, Figure 9, 952 and 942.

Claim 7 – page 12, lines 11-15 Figure 2 218, 220, 222, 224).

Claim 8 – page 12, lines 15-17 , Figure 2, 206, 216, 218.

Claim 9 – page 12, lines 17-18.

Claim 10 – page 9, lines 14-18, Figure 3, 322.

Claim 11 – page 4, lines 23-24, Figure 1 106, 108, 110, 112.

Claim 45 - page 14, lines 13-15; page 13, lines 5-10, Figure 4B, step 434

Claim 46 – page 4, line 25 – page 5, line 1, Figure 9, 952 and 942.

Claim 47 - page 5, lines 1-8.

Claim 48 – page 10, lines 4-10 Figure 3, 302, 304.

Claim 49 – page 22, lines 21-28, Figure 9, 952 and 942.

Claim 50 – page 12, lines 11-15 Figure 2 218, 220, 222, 224).

Claim 51 – page 12, lines 15-17 , Figure 2, 206, 216, 218.

Claim 52 – page 12, lines 17-18.

Claim 53 – page 9, lines 14-18, Figure 3, 322.

Claim 54 – page 4, lines 23-24, Figure 1 106, 108, 110, 112.

## VI GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL (37 C.F.R. 41.37 (c)(1)(vi))

Whether claims 1-11 and 44-54 are patentable under 35 U.S.C. 103(a) over U.S. Patent No. 6,195,703 (Blumenau) in view of U.S. Patent No. 6,157,927 (Schaefer.)

## VII ARGUMENT (37 C.F.R. §41.37(c)(1)(vii))

**Prima facie obviousness has not been established because the combination of Blumenau and Schaefer does not teach or suggest the structure recited in claims 1-11 and 44-54.**

Obviousness is a legal conclusion based on factual evidence. Graham v. John Deere Co., 383 US 1, 148 USPQ 459 (1966). To establish the *prima facie* obviousness of a claimed invention, all of the claim limitations must be taught or suggested by the prior art. In re Wada and Murphy, Appeal 2007-3733, citing In re Ochiai, 71 F.3d 1565, 1572 (Fed. Cir. 1995) and In re Royka, 490 F.2d 981, 180 USPQ 580 (CCPA 1970).

The examiner refers to column 2, lines 5-12 of the Blumenau reference. There the reference discloses a storage system in which a host can send a request to access storage resources through a switch to one of a plurality of network ports of a data storage subsystem. The switch allows paths to be selected between a particular host and alternative storage subsystem ports based on load balancing considerations. However, once a request arrives at any network port of the storage subsystem a data address in the request is translated to a physical location on one of the storage devices and the data is written to, or read from, that storage device. For example, Blumenau, column 4, lines 43-53, discloses:

"If the data to be accessed resides in the cache memory, then the port adapter accesses the data in the cache memory. If the data to be accessed does not reside in the cache memory, then the port adapter forwards a storage access request to the storage adapters 37, 38. One of the storage adapters 37, 38 responds to the storage access request by performing a logical-to-physical translation to determine where the data to be accessed resides on the storage devices, and reads the data from the storage devices and writes the data to the cache memory, for access by the port adapter." (emphasis added)

This section clearly indicates that the physical storage location depends on the data address and not on load balancing considerations. Further sections of Blumenau reinforce this section. For example, the examiner contends that column 12, lines 26-33 of Blumenau disclose a data storage request that includes a data identifier. However, this section of Blumenau actually discusses the contents of entries in a routing table that controls the switch. As stated there, each entry include a list pointer for the host, a storage port list for the host, a host name, a world-wide name for the host, and a source address (S\_ID) for the host. All of this information refers to the host, not to data.

Consequently, although data access requests can be routed along different paths between a host and the data storage subsystem, there is no teaching or suggestion in Blumenau that the data to which the requests refer could be stored in different storage devices based on load balancing considerations. The examiner contends “One [of] ordinary skill in the art at the time of the invention knows that ‘routing information based on the loading characteristic of the storage access request received at the switch inputs in order to balance loading of the storage access requests upon the outputs of the switch’ is prime example of load balancing, where the routing information is based on the loading characteristic of the meaning workload of the memory.” While applicants do not disagree with this contention, it is not what is being claimed. The claims are concerned with where the data is stored not how the data arrives there. For example, claim 1 (and claim 44 which has similar wording) recites:

“... an access interface module which ... based on a workload instead of a physical location in the plurality of devices, dynamically selects a subset of the plurality of storage devices in which the data is stored so that the physical storage device location to which data corresponding to the same data identifier is transferred can change from request to request in order to dynamically distribute the workload across the plurality of storage devices.”

This claim wording does not read onto Blumenau even if the entire Blumenau storage subsystem is considered to be a storage “device” because, under this interpretation, it would not be possible to select a subset of a plurality of storage devices as recited. It is also clear that, while the Blumenau system can alleviate overloading of

the data storage subsystem network ports it cannot accomplish the objectives of the invention as set forth above.

The Schaefer reference discloses a distributed transaction processing system in which there are two different transaction processing environments that operate with different transaction protocols (XATMI and non-XATMI). Two programs, one called a connection manager and the other a resource manager, are used to translate between the two protocols so that an application program operating in one environment can access a resource located in the other environment as noted by the examiner at column 8, lines 25-29. However, the operation of the Schaefer system indicates that it uses conventional database transaction processing in which information is retrieved from, and stored back into, the same database location. Specifically, the storage location is determined by the storage address and not by the workload of the storage resources. The examiner further cites column 9, lines 31-41. However, this section of Schaefer discloses a failure recovery mechanism that uses a common log entry and does not discuss where data is stored.

Therefore, neither of the cited references discloses a system in which data corresponding to the same data identifier or address may be stored in different physical storage device locations based on criteria, such as load balancing or availability and, consequently, the combination of these references cannot teach or suggest this type of operation. Accordingly, claims 1 and 44 patentably distinguish over the cited reference combination.

Claims 2-11 and 45-54 are dependent, either directly or indirectly, on claims 1 and 44, respectively, and incorporate the limitations thereof. Therefore, they also distinguish over the cited reference combination in the same manner as claims 1 and 44. In addition, these claims recite further elements not disclosed or suggested by the cited reference combination.

Claims 2 and 45 recite that the access interface module selects the subset of the plurality of resources based on the relative demand placed on the subset of resources. Claims 1 and 44, from which these claims depend, indicate that this resource subset is where the data is stored. As noted above, neither Blumenau nor Schaefer disclose any arrangement where the data storage location is defined by anything other than the data



address. The examiner cites Schaefer, column 8, lines 42-54 as disclosing this recited element, but this section of Schaefer discusses the incompatibility of the XATMI and non-XATMI protocols and mentions nothing regarding resource selection.

Claims 3, 4, 5, 6, 46, 47, 48 and 49 recite details of the switch fabric that connects the access interface modules and the storage devices. These details include that the switch comprises separate control and data switch fabrics that are optimized for control and data signals, respectively; that the data is transmitted only through the data switch and that the data switch comprises a crosspoint switch and the control switch comprises an Ethernet switch. While it is clear that the Blumenau switch must switch both data requests and the data to which they refer, the switch itself is only shown as a single block (40) with no indication that two separate switches are used. The examiner refers to Blumenau, column 2, lines 7-19. However, this section of Blumenau discusses how data and requests are routed between switch inputs and switch outputs; no internal switch details are provided. Therefore, the elements recited in claims 3 and 46 are not disclosed in Blumenau. Schaefer also discloses no switch analogous to the recitation of these claims.

Claims 7 and 50 further recite a resource module connected to the plurality of resources that generates preallocation information that preallocates services from the plurality of resources in order to evenly distribute a workload across the plurality of resources. As discussed above, Blumenau and Schaefer determine where data will be stored based on the data address; no preallocation is necessary or discussed. The examiner cites Blumenau, column 2, lines 5-21. This section of Blumenau discusses routing of data requests and, accordingly, no preallocation of any resources is discussed.

Claims 8 and 51 recite that the preallocation information is transferred to the access interface module through the switch. The examiner cites Schaefer, column 28, lines 22-37. This section of Schaefer discloses that the resource manager that controls the desired resource receives information regarding a transaction from a transaction object in the form of a cookie. However, the preallocation information is recited in the parent claims 7 and 49 as information that that preallocates services from the plurality of resources in order to evenly distribute a workload across the plurality of resources. It is

clear that the cookie to which Schaefer refers does not do this and, consequently, is not analogous to the recited preallocation information.

Claims 9 and 52 recite that an access interface module selects a subset of the plurality of resources based on the preallocation information. As mentioned previously, the Blumenau and Schaefer references store data in storage devices at location that are determined by addresses, not preallocation information. Therefore, these claims also distinguish over the cited reference combination.

Respectfully submitted

\_\_\_\_\_/paul e. kudirka/\_\_\_\_\_  
Paul E. Kudirka, Esq. Reg. No. 26,931  
LAW OFFICES OF PAUL E. KUDIRKA  
Customer Number 64967  
Tel: (617) 357-0010 Fax: (617) 357-0035

# **VIII APPENDIX OF CLAIMS (37 C.F.R. §41.37(c)(1)(viii))**

The text of the claims involved in the appeal is:

1. Apparatus for providing high-performance, scalable data storage services to a client from a plurality of storage devices, comprising:
  - an access interface module which receives data storage requests, each including a data identifier, and data to be stored from the client and, in response to each service request and based on a workload instead of a physical location in the plurality of devices, dynamically selects a subset of the plurality of storage devices in which the data is stored so that the physical storage device location to which data corresponding to the same data identifier is transferred can change from request to request in order to dynamically distribute the workload across the plurality of storage devices; and
  - a switch fabric for temporarily connecting the access interface module to the selected subset of the plurality of storage devices so that the data can be transferred to the selected subset of storage devices.
2. The apparatus of claim 1 wherein the access interface module selects the subset of the plurality of resources based on the relative demand placed on the subset of resources.

- 1 3. The apparatus of claim 1 wherein the switch fabric comprises a control switch  
2 fabric for transferring control information and a separate data switch fabric for  
3 transferring data.
- 1 4. The apparatus of claim 3 wherein the control switch fabric is optimized for  
2 transferring control information and the data switch fabric is optimized for  
3 transferring data.
- 1 5. The apparatus of claim 3 wherein the request for service includes control  
2 information and data and wherein the access interface module separates the  
3 control information and the data and transfers the data to the selected subset of  
4 resources over the data switch fabric.
- 1 6. The apparatus of claim 3 wherein the data switch fabric comprises a non-  
2 blocking crossbar switch for data transfer and the control switch fabric comprises  
3 an Ethernet switch for control information transfer.
- 1 7. The apparatus of claim 1 further comprising a resource module connected to the  
2 plurality of resources for generating preallocation information that preallocates  
3 services from the plurality of resources in order to evenly distribute a workload  
4 across the plurality of resources.

- 1 8. The apparatus of claim 7 wherein the switch fabric connects the access interface  
2 module to the resource module so that the resource module can transfer the  
3 preallocation information to the access interface module.
- 1 9. The apparatus of claim 8 wherein the access interface module selects a subset  
2 of the plurality of resources based on the preallocation information.
- 1 10. The apparatus of claim 1 wherein the access interface module comprises a data  
2 memory which temporarily stores information transferred between the access  
3 interface module and the selected subset of the plurality of resources.
- 1 11. The apparatus of claim 1 further comprising a plurality of access interface  
2 modules each access interface module receiving service requests from a plurality  
3 of clients.
- 1 44. A method for providing high-performance, scaleable data storage services to a  
2 client from a plurality of storage devices, the method comprising  
3 (a) providing an access interface module which receives data storage  
4 requests, each including a data identifier, and data to be stored from the  
5 client;  
6 (b) using the access interface module in response to each data storage  
7 request and based on a workload instead of a physical location in the  
8 plurality of devices to dynamically select a subset of the plurality of

9 storage devices to which the data is stored so that the physical storage  
10 device location to which data corresponding to the same data identifier is  
11 transferred can change from request to request in order to dynamically  
12 distribute the workload across the plurality of storage devices; and  
13 (c) using a switch fabric to temporarily connect the access interface module to  
14 the selected subset of the plurality of storage devices so that the data can  
15 be transferred to the selected subset of storage devices.

1 45. The method of claim 44 wherein step (b) comprises selecting the subset of the  
2 plurality of resources based on the relative demand placed on the subset of  
3 resources.

1 46. The method of claim 44 wherein step (c) comprises:  
2 (c1) using a control switch fabric for transferring control information; and  
3 (c2) using a separate data switch fabric for transferring data.

1 47. The method of claim 46 wherein step (c1) comprises optimizing the control  
2 switch fabric for transferring control information and step (c2) comprises  
3 optimizing the data switch fabric for transferring data.

1 48. The method of claim 46 wherein the request for service includes control  
2 information and data and wherein step (b) comprises separating the control

information and the data and step (c) comprises transferring the data to the selected subset of resources over the data switch fabric.

49. The method of claim 46 wherein step (c1) comprises using a non-blocking crossbar switch for data transfer and step (c2) comprises using an Ethernet switch for control information transfer.

50. The method of claim 44 further comprising:

(d) providing a resource module connected to the plurality of resources; and

(e) using the resource module to generate preallocation information that preallocates services from the plurality of resources in order to evenly distribute a workload across the plurality of resources.

51. The method of claim 50 wherein step (c) comprises connecting the access interface module to the resource module so that the resource module can transfer the preallocation information to the access interface module.

52. The method of claim 51 wherein step (b) comprises selecting a subset of the plurality of resources based on the preallocation information.

53. The method of claim 44 wherein step (b) comprises temporarily storing information transferred between the access interface module and the selected subset of the plurality of resources.

1    54.    The method of claim 44 wherein step (a) further comprises providing a plurality of  
2        access interface modules each access interface module receiving service  
3        requests from a plurality of clients.



**IX EVIDENCE APPENDIX (37 C.F.R. §41.37(c)(1)(ix))**

**None**

**X RELATED PROCEEDINGS APPENDIX (37 C.F.R. §41.37(c)(1)(x))**

**NONE**